Martín-Vázquez, R., Aler, R. y Galván, I. M. (2018). Wind Energy Forecasting at Different Time Horizons with Individual and Global Models. In *Artificial Intelligence Applications and Innovations (AIAI 2018)*, 519, pp. 240-248.

DOI: https://doi.org/10.1007/978-3-319-92007-8_21

Wind Energy Forecasting at Different Time Horizons with Individual and Global Models

R. Martín-Vázquez(✉), R. Aler, and I. M. Galván

EVANNAI, Carlos III University of Madrid,
Avda. Universidad, 30, 28911 Leganés, Spain
ruben.martin@uc3m.es

Abstract. In this work two different machine learning approaches have been studied to predict wind power for different time horizons: individual and global models. The individual approach constructs a model for each horizon while the global approach obtains a single model that can be used for all horizons. Both approaches have advantages and disadvantages. Each individual model is trained with data pertaining to a single horizon, thus it can be specific for that horizon, but can use fewer data for training than the global model, which is constructed with data belonging to all horizons. Support Vector Machines have been used for constructing the individual and global models. This study has been tested on energy production data obtained from the Sotavento wind farm and meteorological data from the European Centre for Medium-Range Weather Forecasts, for a 5 × 5 grid around Sotavento. Also, given the large amount of variables involved, a feature selection algorithm (Sequential Forward Selection) has been used in order to improve the performance of the models. Experimental results show that the global model is more accurate than the individual ones, specially when feature selection is used.

Keywords: Wind power forecasting · Machine learning · Forecasting horizons

1 Introduction

The correct forecast of the energy obtained from the wind is still one of the main challenges of renewable energies. To predict the wind power generated in a wind farm is a difficult task although there is a wide literature on this area. Currently, to forecast the power there are mainly three approaches: physical methods, statistical methods and methods based on artificial intelligence [1, 2].

Physical methods collect values measured in the lower atmosphere to build mathematical models, named Numerical Weather Prediction (NWP), that are used to make predictions of the weather [3]. These methods have the drawback that they require long operation time and a large amount of computational resources.

Statistical methods use historical data from wind farms to find a relationship between the input variables and the power. These models are faster than physical methods but they have the disadvantage that they have large error for long horizons. Time series [4], Autoregressive Moving Average (ARMA) [5], Autoregressive

Integrated Moving Average (ARIMA) [6] or Modified Taylor Kriging (MTK) [7], are examples of these techniques.

Artificial intelligence methods use mainly, but not exclusively, machine learning techniques. They use information from NWP variables and/or historical data to predict power wind, trying to find the relationship between these variables. Some of these artificial intelligence techniques are Fuzzy Logic [8], Genetic Algorithms [9], Artificial Neural Networks (ANN) [10], Support Vector Machines (SVM) [11] or ensemble methods [12].

Some of these approaches can be hybridized. The growing popularity of renewable energies has led to the search for new alternatives by combining existing methods to try to make more reliable predictions. Time series and boosting with historical data [13], ANN and time series with wind speed [14] and Gaussian Process with NWP and historical data [15] are some of the combinations that can be found.

In this work, machine learning methods using Support Vector Machine models (SVM) are proposed to predict wind power at different time horizons (or steps), from 3 h until 15 h, using meteorological variables (from a NWP model) in a grid centered at the Sotavento wind farm. With this purpose two approaches has been studied: individual models and global models. Individual models are made for working in a single time horizon, building up a different model for each horizon. The motivation for using individual models is to train them under the same situation that they will be used operationally. When used operationally, the only inputs available to the model are the meteorological variables forecasts. Therefore, for training the individual models, forecasts carried out in the past (historical forecasts) will also be used. This has the advantage that both training and operation use the same kind of data, but it has two drawbacks. First, forecasts contain prediction errors that might affect negatively the training process. Second, each individual model can only be trained with data belonging to its time horizon. Thus, a second approach is proposed, the global model, where a single model is trained using the data from all horizons put together. In this case, the model can be trained using reanalysis data, that contains observations/ measurements, rather than forecast records. In operation mode, the same global model can be used for making predictions at any horizon.

On the other hand, using a grid of meteorological variables implies that the number of attributes can be very large, which can reduce the model generalization capability, specially for the individual models, which are trained with fewer data. In a previous work [16] three different attribute selection methods were evaluated to estimate wind energy at Sotavento. It was concluded that Sequential Forward Selection (SFS) reduced significantly the number of attributes while obtaining very accurate results. In this work, SFS has also been used for both the individual and global models. In the case of the individual models, different attributes can be selected for different horizons.

The rest of the paper is organized as follows. In Sect. 2 data sets are explained. Section 3 describes the individual and global approaches that are studied in the present work. Experimental results are presented in Sect. 4. Finally, Sect. 5 draws the conclusions learned from this study.

2 Data

As in the previous work [16], data is obtained from a 5 × 5 grid, with a distance between coordinates of 0.125 degrees, centered at Sotavento experimental wind farm between 2005 and 2010. 19 meteorological variables are selected for each grid point, and they have been obtained from the ERA-20C dataset made available by the European Centre for Medium-Range Weather Forecasts (ECMWF)[1].

This results in a total of 475 attributes (19 * 5 * 5). The 19 meteorological variables are shown in Table 1. For the same period, there is a reanalysis data set and a forecasting data set. As its name indicates, forecasting datasets contain forecasts, while reanalysis datasets contain information obtained to observations. The forecast datasets provide predictions at different steps (or horizons). All the forecasts are made every day at 06:00UTC and predictions are provided for steps 3, 6, 9, 12, and 15 h ahead. The reanalysis dataset contains a time series for each meteorological variable, with observations provided at 06:00UTC+3h, 06:00UTC+6h … 06:00UTC+15h. Generated wind power has been obtained from the Sotavento website[2]. Table 2 displays the number of instances for each data set.

3 Individual and Global Approaches for Energy Forecasting

In this work, two different approaches are used for predicting wind energy: individual models and a global model. Individual models are built to be used for specific time horizons while the global model can be used for any time horizon. Next, the process used for training and using the individual and global models will be described:

Individual models follow Eq. 1, where t_0 is the time at which the predictions are made (06:00UTC), h = 3, 6, 9, 12, 15 are the steps, and $v_i(t_0 + h)$ are the forecasted meteorological variables. Each F_h model is trained using the data for that particular step from the forecast dataset (see Table 2, so that every individual model is specialized in a particular step.

$$E_h^{(0)}(t_0+h) = F_h(\widehat{v_1}(t_0+h), \ldots, \widehat{v_n}(t_0+h)) \quad (1)$$

In order to study the influence on the forecast of the energy measured at prediction time t_0, a new input variable is added (see Eq. 2) with the energy value at 06:00 UTC ($E(t_0)$).

$$E_h^{(1)}(t_0+h) = F_h(\widehat{v_1}(t_0+h), \ldots, \widehat{v_n}(t_0+h), E(t_0)) \quad (2)$$

The **global model** is trained using reanalysis data, according to Eq. 3, where $v_i(t)$ are the reanalysis meteorological variables. In this case, the data size is larger than the ones used with individual models (see Table 2).

[1] https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-20c.

[2] http://www.sotaventogalicia.com/en/real-time-data/historical.

Table 1. Meteorological variables used as input to the models.

Variables	
2 m temperature (K)	10 m U wind component (ms^{-1})
10 m V wind component (ms^{-1})	100 m U wind component (ms^{-1})
100 m V wind component (ms^{-1})	Convective available potential energy (Jkg^{-1})
Forecast logarithm of surface roughness for heat	Forecast surface roughness (m)
Instantaneous eastward turbulent surface stress (Nm^{-2})	Instantaneous northward turbulent surface stress (Nm^{-2})
Leaf area index, high vegetation (m^2m^{-2})	Leaf area index, low vegetation (m^2m^{-2})
Neutral wind at 10 m u-component (ms^{-1})	Neutral wind at 10 m v-component (ms^{-1})
Soil temperature level 1 (K)	Soil temperature level 2 (K)
Soil temperature level 3 (K)	Soil temperature level 4 (K)
Surface pressure (Pa)	

Table 2. Number of instances for each data set.

Data Set	Instances
Forecast step 3	1445
Forecast step 6	1446
Forecast step 9	1543
Forecast step 12	1550
Forecast step 15	1468
Reanalysis	8347

$$E^{(0)}(t) = F(v_1(t), \ldots, v_n(t)) \quad (3)$$

Equation 4 shows how to use the global model for making predictions at time t_0 for different horizons (h = 3, 6, 9, 12, 15). In this case, forecasts $\widehat{v_i}(t_0 + h)$ must be used because observations are not available at prediction time. It is important to remark that, differently to individual models, the same global model is used for all steps.

$$E^{(0)}(t_0 + h) = F(\widehat{v_1}(t_0 + h), \ldots, \widehat{v_n}(t_0 + h)) \quad (4)$$

Similarly to individual models, the influence of wind energy measured at prediction time ($E(t_0)$) is studied, with models trained according to Eq. 5.

$$E^{(1)}(t) = F(v_1(t), \ldots, v_n(t), E(t_0)) \quad (5)$$

Similarly to model $E^{(0)}$, in order to use $E^{(1)}$ for making predictions, Eq. 6 must be used.

$$E^{(1)}(t_0+h) = F(\widehat{v_n}(t_0+h), \ldots, \widehat{v_n}(t_0+h), E(t_0)) \tag{6}$$

4 Experimental Results

In this section the results obtained in the experiments are shown. To measure the performance the Mean Absolute Error (MAE) is used. A methodology similar to cross validation has been used to evaluate both individual and global models. Given that data is available for six years (2005 to 2010), each year is used as a test set, and the remaining five years are used to construct the model. That means that the process of model construction and model evaluation has been repeated six times (6 folds) and then the average number of inputs and test MAE have been obtained. Data reserved for model construction is further divided into training and validation datasets by choosing, every three days, the first two days for training and the third one for validation. The validation partition will be used for hyper-parameter tuning and to select the most relevant attributes with the SFS algorithm. The final model is trained joining the training and validation data, using the hyper-parameters and attributes selected previ-ously. This model is finally evaluated on the test set.

SVM hyper-parameter tuning has been carried out by means of grid search, for Gamma = {0, 0.01, 0.1, 0.1, 0.25, 0.5, 0.75, 1} and C = {0.1, 1, 10, 100, 500}. Grid search returned the same parameter values for all folds: C = 1 and Gamma = 0.01 for individuals and global models.

Because of the large number of input attributes, and taking into account the results of the previous work, SFS has been used to reduce the number of input features. In the case of the global model, the 26 attributes selected are the same as in the previous work, where also the Sotavento reanalysis dataset was used [16]. In the case of the individual models, SFS was run for each of the steps, so that the attributes selected are specialized for every step.

The average results of the 6 folds are shown in Table 3. The first half of the table displays results for the individual models and the second half for the global one. For each of them, the MAE is provided for the model without feature selection (second column), with SFS selection (third column), and with SFS selection plus the $E(t_0)$ input (last column). All results have been broken down by step (first column), but average and standard deviation results are also included. When SFS selection is used, the average number of inputs selected is also given.

With respect to the individual models, it can be seen that using SFS (third column) reduces significantly the number of inputs (from 475 to 8.54 on average). MAE is also improved for all steps except the last one. Adding $E(t_0)$ as input to the model ($E_h^{(1)}$ in the fourth column) does not improve results, except for the first step (3 h) (see Fig. 1). For the global model, SFS selection improves MAE for all steps and adding $E(t_0)$ as input decreases MAE for the two first steps (see Fig. 2). This should be expected, because the longer the horizon, the less relevant is the power measured at prediction time t_0.

Table 3. Results for the individual global approaches.

Individual models				
Step	$E_h^{(0)}$, 475 Inputs	$E_h^{(0)}$, SFS Inputs		$E_h^{(1)}$, SFS Inputs
	MAE	MAE	Number Inputs	MAE
Step 3	266.15	261.34	7.33	246.79
Step 6	291.42	272	11.17	273.34
Step 9	280.06	275.32	11.17	306.81
Step 12	259.32	257.5	6.67	295.41
Step 15	288.64	303.21	6	356.82
Average	277.12	274.07	8.47	295.83
Std.	13.99	17.94	2.51	41.08
Global models				
Step	$E^{(0)}$, 475 Inputs	$E^{(0)}$, SFS Inputs		$E^{(1)}$, SFS Inputs
	MAE	MAE	Number Inputs	MAE
Step 3	268.86	256.61	26	238.8
Step 6	282.88	259.74	26	251.3
Step 9	273.42	247.71	26	252.46
Step 12	260.53	237.13	26	242.6
Step 15	295.72	281.33	26	303.43
Average	276.28	256.5	26	257.72
Std.	13.54	16.43	0	26.2

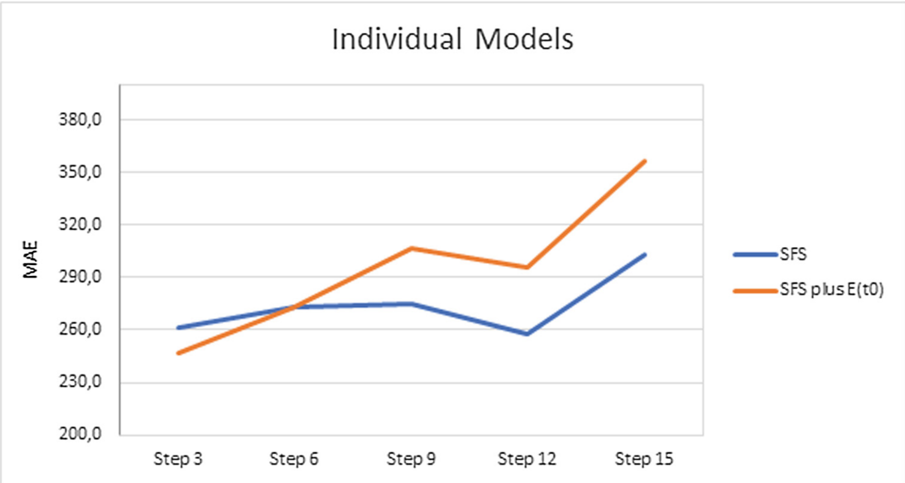

Fig. 1. Individual models with/without $E(t_0)$

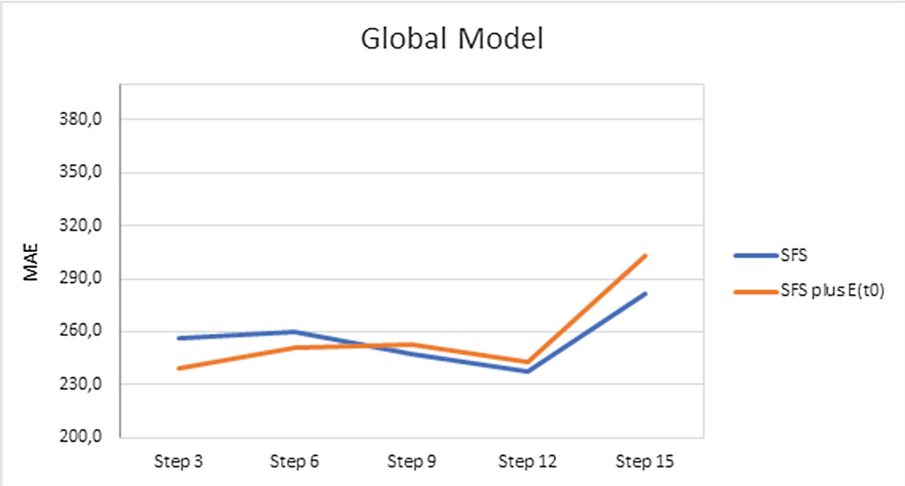

Fig. 2. Global model with/without $E(t_0)$

Comparing the individual and global models, it can be seen in Table 3 that the global models outperform the individual ones on average in the three cases studied: without feature selection, with SFS, and with SFS plus $E(t_0)$. In the first case, the improvement of the global model is small (277.12 vs. 276.28), but the difference increases for the other two cases (274.07 vs. 256.50 and 295.83 vs. 257.72, respec-tively). In fact, for the last two cases, the global model is more accurate than the individual ones, not only on average, but for each step.

In order to assess the significance of the comparisons between global and individual models, the number of folds where each model is better will be given. $E^{(0)}$ global model is better in 5 out of 6 folds for every step than individual models. For $E^{(1)}$ models with SFS variables, results are better for global model than individual models in step 3 for 5 folds. In the following steps, the global model is better than the individual models in all of the folds.

In order to provide some baseline results, the persistence method has also been used. Persistence uses the current value of wind power as forecast for the next time step. The 6-fold average MAE obtained for each time horizon can be seen in Table 4.

Table 4. Results for the persistence method.

Step	MAE
Step 3	305.85
Step 6	395.20
Step 9	427.57
Step 12	450.34
Step 15	494.20
Average	414.63
Std.	25.8

It can be observed that the different approaches studied in this work are better than persistence for all the steps.

5 Conclusions

Two different approaches, individual and global, for predicting wind power in different time horizons have been studied and compared. Individual models are specific for each step. Therefore, they are trained with data belonging to the step and with forecasting dataset. The global model is trained with data from all steps and using the reanalysis dataset. This results in a unique model that can be used for making predictions in the different steps.

Although it could be expected that individual models performed better because they are specific for each horizon, the fact that fewer data is used to train each of them, and the need to use the forecasted input meteorological variables, rather than the reanalysis dataset, makes the global model perform more accurately, at least in this study.

Regarding the use of the SFS algorithm for feature selection, results show that both individual and global improve accuracy, but it is the global approach which specially takes advantage from it. Finally, adding the wind power measurement available at prediction time as input to the models, benefits mainly the global model and only the forecasting horizons close to prediction time. The improvement for short time horizons (steps 3 and 6) should be expected, because long time horizons depend much less on observations at prediction time.

Acknowledgements. The authors acknowledge financial support granted by the Spanish Ministry of Science under contract ENE2014-56126-C2-2-R.

References

1. Tascikaraoglu, A., Uzunoglu, M.: A review of combined approaches for prediction of short-term wind speed and power. Renew. Sustain. Energy Rev. **34**(Suppl. C), 243–254 (2014)
2. Jung, J., Broadwater, R.P.: Current status and future advances for wind speed and power forecasting. Renew. Sustain. Energy Rev. **31**, 762–777 (2014)
3. Roulston, M.S., Kaplan, D.T., Hardenberg, J., Smith, L.A.: Using medium-range weather forcasts to improve the value of wind energy production. Renew. Energy **28**(4), 585–602 (2003)
4. Croonenbroeck, C., Ambach, D.: A selection of time series models for short- to medium-term wind power forecasting. J. Wind Eng. Ind. Aerodyn. **136**, 201–210 (2015)
5. Liu, H., Erdem, E., Shi, J.: Comprehensive evaluation of ARMA-GARCH(-M) approaches for modeling the mean and volatility of wind speed. Appl. Energy **88**, 724–732 (2011)
6. Kavasseri, R.G., Seetharaman, K.: Day-ahead wind speed forecasting using f-ARIMA models. Renew. Energy **34**(5), 1388–1393 (2009)
7. Liu, H., Shi, J., Erdem, E.: Prediction of wind speed time series using modified Taylor Kriging method. Energy **35**(12), 4870–4879 (2010)

8. Zhu, B., Chen, M., Wade, N., Ran, L.: A prediction model for wind farm power generation based on fuzzy modeling. Procedia Environ. Sci. **12**, 122–129 (2012)
9. Damousis, I.G., Alexiadis, M.C., Theocharis, J.B., Dokopoulos, P.S.: A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. IEEE Trans. Energy Convers. **19**(2), 352–361 (2004)
10. Mohandes, M.A., Halawani, T.O., Rehman, S., Hussain, A.A.: Support vector machines for wind speed prediction. Renew. Energy **29**(6), 939–947 (2004)
11. Heinermann, J., Kramer, O.: Precise wind power prediction with SVM ensemble regression. In: Wermter, S., Weber, C., Duch, W., Honkela, T., Koprinkova-Hristova, P., Magg, S., Palm, G., Villa, Alessandro E.P. (eds.) ICANN 2014. LNCS, vol. 8681, pp. 797–804. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11179-7_100
12. Ren, Y., Suganthan, P.N., Srikanth, N.: Ensemble methods for wind and solar power forecasting—a state-of-the-art review. Renew. Sustain. Energy Rev. **50**, 82–91 (2015)
13. Jiang, Y., Chen, X., Yu, K., Liao, Y.: Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm. J. Mod. Power Syst. Clean Energy **5**(1), 126–133 (2017)
14. Doucoure, B., Agbossou, K., Cardenas, A.: Time series prediction using artificial wavelet neural network and multi-resolution analysis: application to wind speed data. Renew. Energy **92**, 202–211 (2016)
15. Chen, N., Qian, Z., Nabney, I.T., Meng, X.: Wind power forecasts using Gaussian processes and numerical weather prediction. IEEE Trans. Power Syst. **29**(2), 656–665 (2014)
16. Martín-Vázquez, R., Aler, R., Galván, Inés M.: A study on feature selection methods for wind energy prediction. In: Rojas, I., Joya, G., Catala, A. (eds.) IWANN 2017. LNCS, vol. 10305, pp. 698–707. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59153-7_60